

Protein domain screening system for structural genomics

○Eiko Seki¹, Takanori Kigawa^{1,2}, Natsuko Matsuda¹, Yoshihide Hayashizaki³, Shigeyuki Yokoyama^{1,2,4}

¹*Protein Research G., GSC, RIKEN Yokohama Institute*, ²*Cell Signal Lab., RIKEN Harima Institute at Spring-8*, ³*Genome Exploration Research G., GSC, RIKEN Yokohama Institute*, ⁴*Grad. Sch. of Sci., Univ. of Tokyo*

In RIKEN GSC Protein Research Group, we are heading for the systematical determination of the three-dimensional structures of protein folds. For this purpose, it is essential to prepare soluble and properly folded protein samples for structural analyses.

Here, we constructed a new experimental system, which enables us to select suitable samples for structural analyses. To begin with, target cDNA was randomly fragmented with Exonuclease III, to construct a deletion library of various lengths. For the 1st step screening, deleted fragments were expressed as C-terminal GFP-fusions in *E.coli*, and fluorescing clones were selected as candidates. For the 2nd step screening, those selected clones were expressed in cell free synthesis system, and screened on the basis of the intensity of GFP fluorescence.

We verified this screening system with mouse growth factor receptor-binding protein 2 (Grb2). The GFP fluorescence in the 2nd step screening was correlated with the solubility of deleted fragments. The terminals of soluble fragments coincided with the structural boundaries of Grb2. Thus, by using this system, we could screen protein samples for structural analyses, from the pool of various proteins. We are now applying this system to define structural domains of mouse cDNA.

序論

構造ゲノム科学は、タンパク質の立体構造解析を系統的かつ網羅的に行うものである。数万個といわれるタンパク質の立体構造は、数千種類の基本構造の組み合わせから形成されていると考えられており、我々はタンパク質の基本構造を網羅的に解明することにより、タンパク質の立体構造の全貌を明らかにしようとしている。この目的のためには、NMRやX線結晶構造解析を用いた立体構造解析に適した、構造をとり、かつ可溶性の高いタンパク質試料を、迅速に選び出す系の構築が重要である。

Waldo (1999) らの報告によると、タンパク質のC末端側にGFPを融合させた状態で大腸菌内で発現させることにより、その可溶性を検出することが可能である。しかしながら、大腸菌内で発現したタンパク質の蛍光強度を測定することは、煩雑であり、多検体同時処理に向いているとはいえない。そこで、我々は、cDNA deletion libraryのGFP-fusionと無細胞タンパク質合成系を組み合わせることにより、より正確かつ簡便に、可溶性のタンパク質ドメインを選び出す系の構築を行った。

タンパク質ドメイン選択法の開発

- 実験的タンパク質ドメイン選択法の必要性
 - 立体構造解析において、構造をとり、かつ可溶なタンパク質試料が必要である。
 - タンパク質の可溶性は、一次配列上からは予測出来ず、実験的手法でしか確認できない。
 - 構造ゲノム科学では多検体同時処理による構造決定が必要である。
- 必須条件
 - 簡便かつ迅速な方法
 - 立体構造をとったドメインの選択
 - 可溶性ドメインの選択
 - 大量合成系への簡便な移行

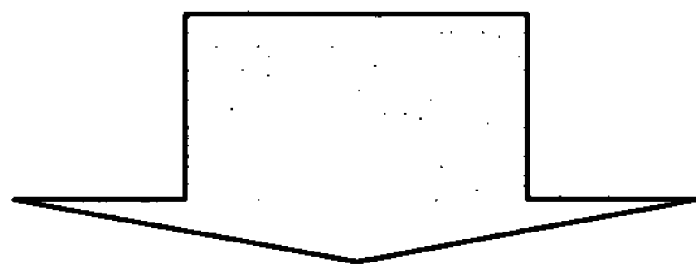
GFP融合体 を利用した 可溶性サンプルのスクリーニング

GFP 融合体:

可溶性タンパク質をC末端側でGFP融合体にすると発光する

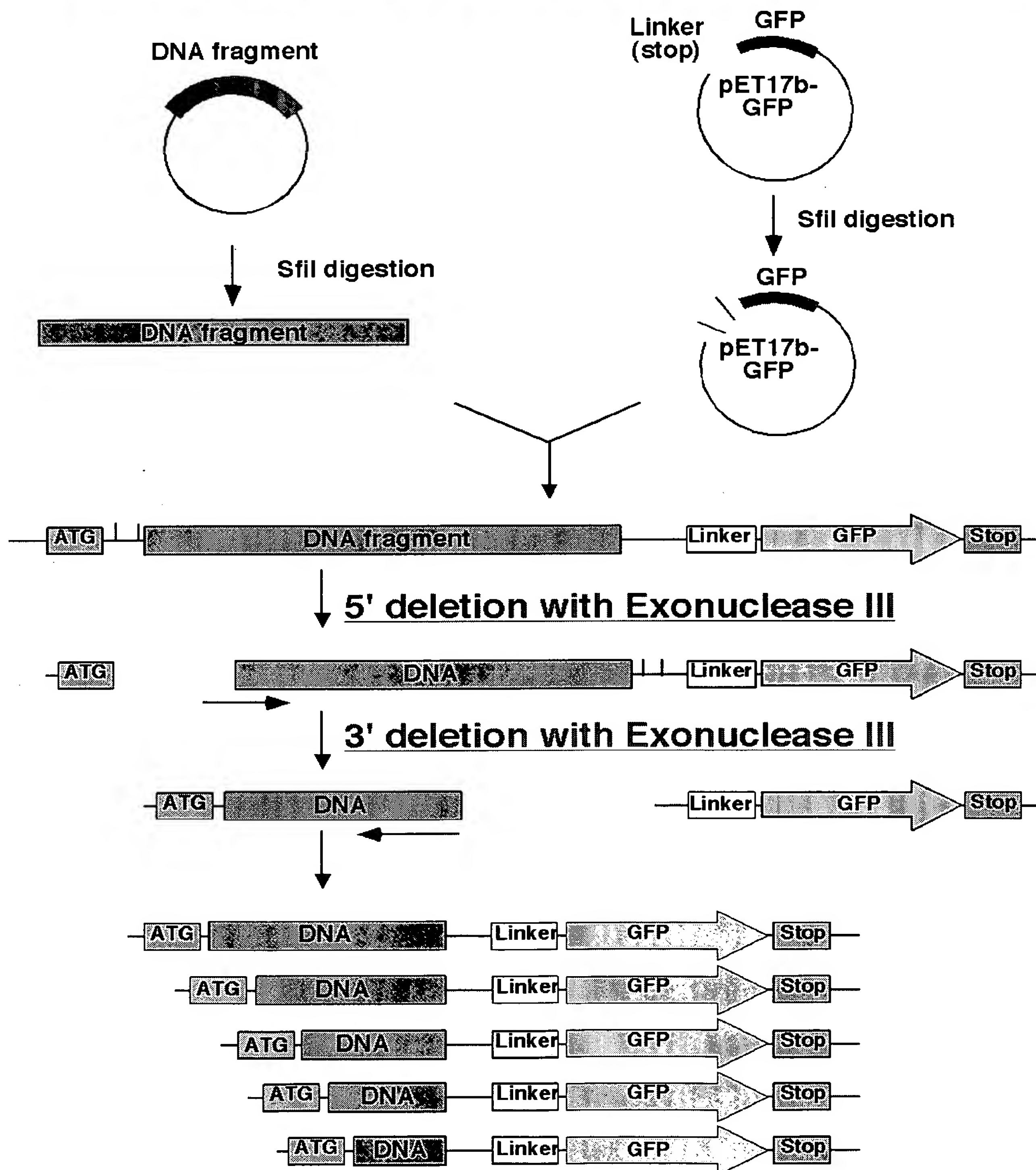
⇔ 難溶性タンパク質は発光しない

(Waldq G S et al Nat Biotechnol 17, 691-695 (1999))



様々な長さを含む均質で良質なdeletion libraryと組み合わせる
ことで、効率よく構造解析に適したタンパク質ドメインをスクリー
ニングできる

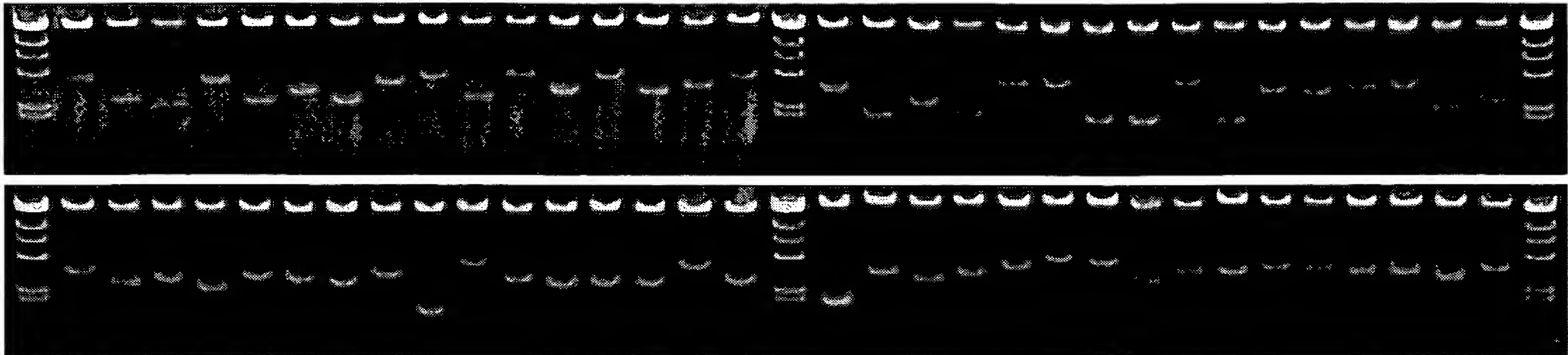
Exonuclease IIIを用いた Deletion Libraryの作成方法



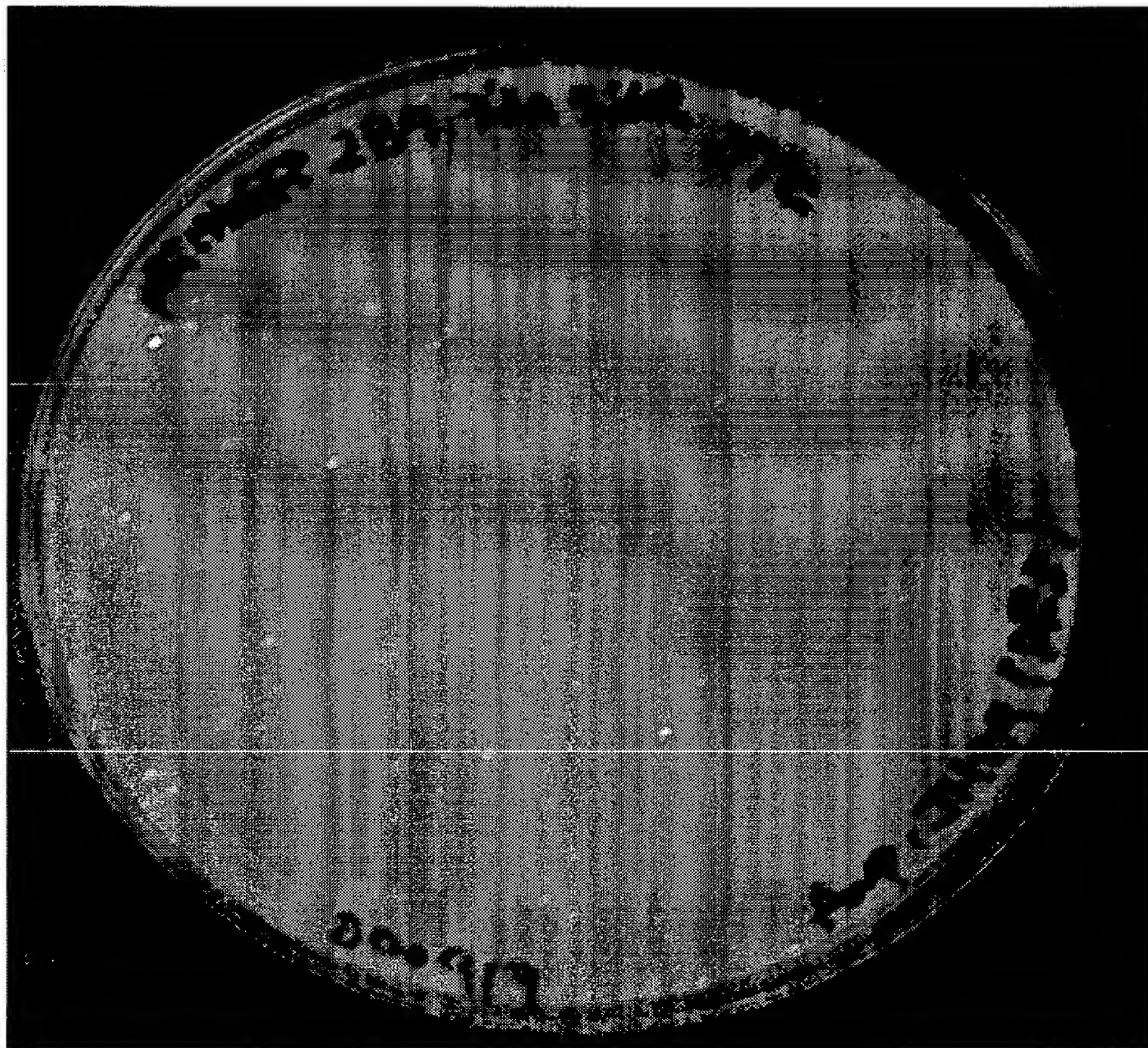
Target DNAの制限酵素サイト検索

Trial	OK				2Kenz change				No enz.	O: 2sides deletion				Δ: 1 side deletion				X: no deletion												
					General ID	P.P rio	MW [Da]	PCR成 功度	R.T 2	R.SUP 2	PstI cut	EcoR I cut	SmaI cut	ApaI	KpnI cut	SpeI cut	NheI cut	3' SphI	3'Nsi I	5'Cla I	5'KaeI	EcoRV	NruI	PmlI	StuI	SacII	NarI	HindIII	NotI	XbaI
	X	14	1B2	76	ZX00001G13 A		10740	3	498.9	411	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
	Δ	161	2F5	70	ZX00003K23 A		16300	3	157.8	94.87	214	###	###	###	###	###	###	###	292	###	###	###	###	###	###	###	###	###	41	###
	X	2	1A2	66	ZX00001E07 A		17121	3	360.9	245.8	238	1	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	19	###
	X	98	1H12	36	ZX00002H05 A		17206	3	187.8	152.3	1331	587	1071	###	208	###	###	1638	###	###	###	###	###	###	25	###	###	###	###	###
	Δ	116	2B8	1	ZX00003M12 A		20930	3	325.1	119.3	310	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	662
	X	70	1F10	38	ZX00002K10 A		22088	3	534.1	96.9	258	###	###	###	###	###	###	600	###	###	###	###	###	###	###	###	###	###	1309	894
	X	38	1D2	59	ZX00001I07 A		25649	3	424	303.1	1404	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
000201	O/Δ	144	2D12	64	ZX00003L10 A		29655	3	440.2	129.5	53	###	###	###	61	###	###	###	###	###	496	###	###	###	###	###	###	41	###	###
	O	56	1E8	48	ZX00002E09 A		32885	3	215.6	202.8	###	7	###	###	###	###	###	###	###	###	###	###	###	###	282	###	###	915	###	###
	O/Δ	46	1D10	40	ZX00002G14 A		33400	3	366.7	131.6	326	###	###	###	###	###	###	###	652	###	###	###	###	###	###	###	###	###	###	###
	O	187	2H7	23	ZX00003M04 A		35674	3	288	216.5	706	###	452	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
	O/Δ	191	2H11	62	ZX00003H22 A		38197	3	25.68	18.69	###	###	###	###	398	###	###	###	###	###	###	###	###	###	302	###	###	1295	###	###
	O	123	2C3	29	ZX00002E09 A		39769	3	193.1	133.3	915	359	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	37	###	###
000201	O	184	2H4	72	ZX00003G09 A		42083	3	57.88	50.12	###	25	1472	###	###	###	###	288	909	1725	###	###	###	###	###	###	###	1177	###	###
000201	O	192	2H12	71	ZX00003F02 A		43663	3	236.9	51.23	222	###	###	783	###	###	###	###	20	###	703	403	###	###	###	###	703	1303	###	537
	O	113	2G5	54	ZX00003M07 A		44603	3	143.3	13.4	502	###	472	###	###	###	###	###	###	###	###	###	###	567	###	###	217	###	1014	
	O/Δ	178	2B11	42	ZX00003N21 A		47702	3	134.5	87.5	###	###	###	138	###	###	###	999	###	###	###	###	###	###	1075	###	###	1450	###	812
000201	O	73	1G1	21	ZX00001C13 A		48119	3	390.1	320.1	635	###	###	659	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	96
	O	62	1F2	7	ZX00001I009 A		50436	3	303.9	137.9	725	###	###	1221	###	###	###	533	983	1255	61	###	###	###	327	###	61	524	29	###
	O	181	2H1	27	ZX00002P05 A		52085	3	26.74	23.77	76	###	79	76	###	###	336	937	###	###	###	###	###	###	###	###	###	679	###	###
	O/Δ	143	2D11	77	ZX00003G12 A		52509	3	152.4	19.1	1444	###	10	###	###	###	###	245	###	###	###	258	###	###	###	###	###	1456	###	###
	O/Δ	75	1G3	32	ZX00001C08 A		54264	3	440.2	384.4	140	###	612	89	###	###	951	579	###	###	###	###	###	###	932	###	1970	###	###	###
	O/Δ	7	1A7	47	ZX00001F14 A		58933	3	482.1	77.87	1391	70	1193	###	941	###	###	###	###	###	###	###	###	###	###	###	###	1075	###	###
	O	166	2F10	24	ZX00003L13 A		61309	3	111.2	96.99	53	525	###	290	230	###	87	###	1421	###	###	###	###	###	978	###	###	1466	###	###
	O	111	2B5	74	ZX00002J24 A		64780	3	29	14.55	694	###	1055	###	56	582	1570	1140	###	###	###	###	###	###	1322	###	###	705	###	###
000201	O	133	2D1	9	ZX00002J21 A		66572	3	67.46	46.15	2165	2676	###	###	1829	###	###	###	1735	###	###	583	###	2232	###	###	###	1927	###	###
	O	78	1F4	4	ZX00001B03 A		69242	3	294.5	190.7	1153	1156	###	###	###	120	###	###	1778	###	###	###	###	###	###	27	###	1142	###	###
000201	O	112	2B4	14	ZX00003A21 A		82257	3	158.1	123.2	6	###	###	###	1804	###	###	###	2562	###	###	###	###	###	865	###	###	###	###	1635
	O/Δ	27	1C3	73	ZX00001M0E A		20905	2	43.03	26.17	158	###	416	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	1002
	X	67	1F7	44	ZX00001L04 A		30333	2	268.3	171.9	838	###	###	###	###	###	324	###	###	###	###	###	###	###	###	###	###	1666	###	2406
	X	69	1F9	28	ZX00002I15 A		31739	2	98.59	50.79	###	###	###	###	###	###	###	###	###	535	###	554	###	###	###	###	###	1279	###	###
	O	53	1E5	53	ZX00001J05 A		49239	2	232.5	67.51	397	###	###	###	###	###	###	###	###	997	###	###	###	###	###	###	###	74	###	###
	Δ	92	1H8	69	ZX00002E23 A		49956	2	188.5	163.4	588	###	447	1326	###	###	###	1424	###	###	###	###	###	###	333	###	###	###	###	###
	O	21	1B9	20	ZX00002G11 A		52936	2	85.18	139.9	229	###	1064	23	159	###	###	###	###	###	###	###	###	692	###	18	###	###	###	###
	O/Δ	88	1H4	57	ZX00001B21 A		71175	2	81.99	60.08	155	###	1052	78	###	###	###	1870	###	###	###	###	###	###	96	###	###	1258	###	1455
	O	1	1A1	56	ZX00001A05 A		76132	2	107.3	62.98	56	###	1205	230	###	###	1698	###	###	1918	1395	###	###	733	813	1053	1395	###	###	###
	X	129	2C9	79	ZX00003F19 A		13178	1	304.5	124.1	380	###	###	###	15	###	###	###	110	###	###	###	###	###	###	###	###	###	###	###
	X	183	2H3	55	ZX00002P10 A		15051	1	280.9	68.85	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	546
	X	176	2G8	45	ZX00003F11 A		24914	1	335.5	118.5	722	58	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###	###
	O/Δ	137	2D5	39	ZX00003I15 A		40009	1	216.5	136.3	509	###	928	###	###	###	###	###	661	###	###	###	###	###	###	###	###	1138	###	###
	O	177	2G9	49	ZX00003J09 A		45841	1	188.5	175.7	###	###	1058	336	###	###	###	515	###	###	687	###	###	###	128	###	687	###	1648	###
	O	124	2C4	52	ZX00003C03 A		46051	1	216.9	137.6	621	###	###	###	###	###	###	396	###	1117	###	886	###	###	###	###	###	804	###	###
	O	61	1F1	46	ZX00001C01 A		51615	1	554.6	260.4	408	###	###	###	###	###	###	###	###	###	###	###	###	374	640	###	###	###	###	###
	O	140	2D8	25	ZX00003D20 A		54874	1	390.4	188.1	###	817	###	###	###	###	###	###	###	###	529	###	###	###	###	###	529	366	###	###
	O	190	2H10	13	ZX00003N09 A		61565	1	303.7	115.5	875	207	1535	###	###	###	###	1098	###	49	###	1201	###	###	1299	###	###	###	1455	###

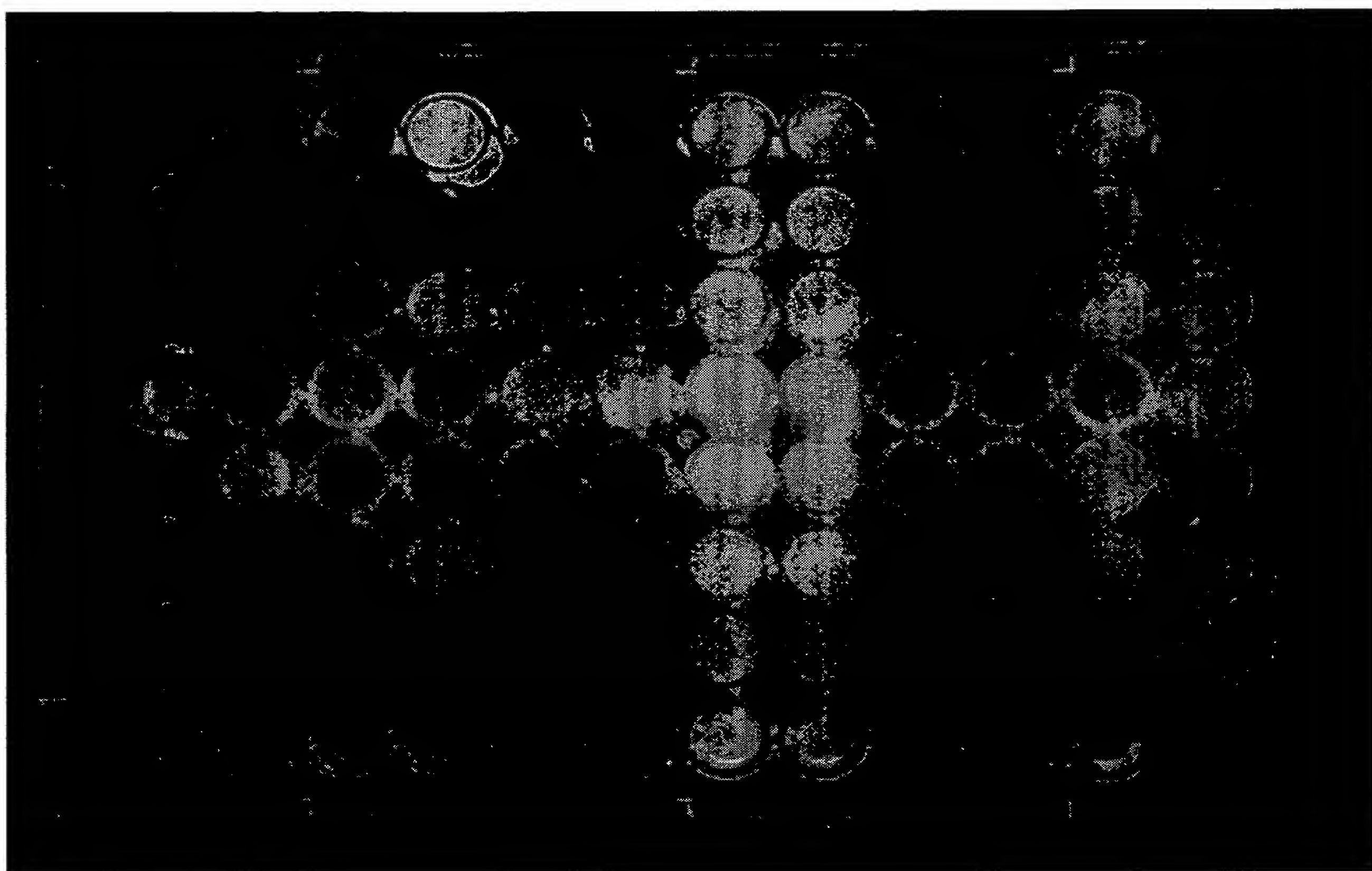
作成したDeletion Library



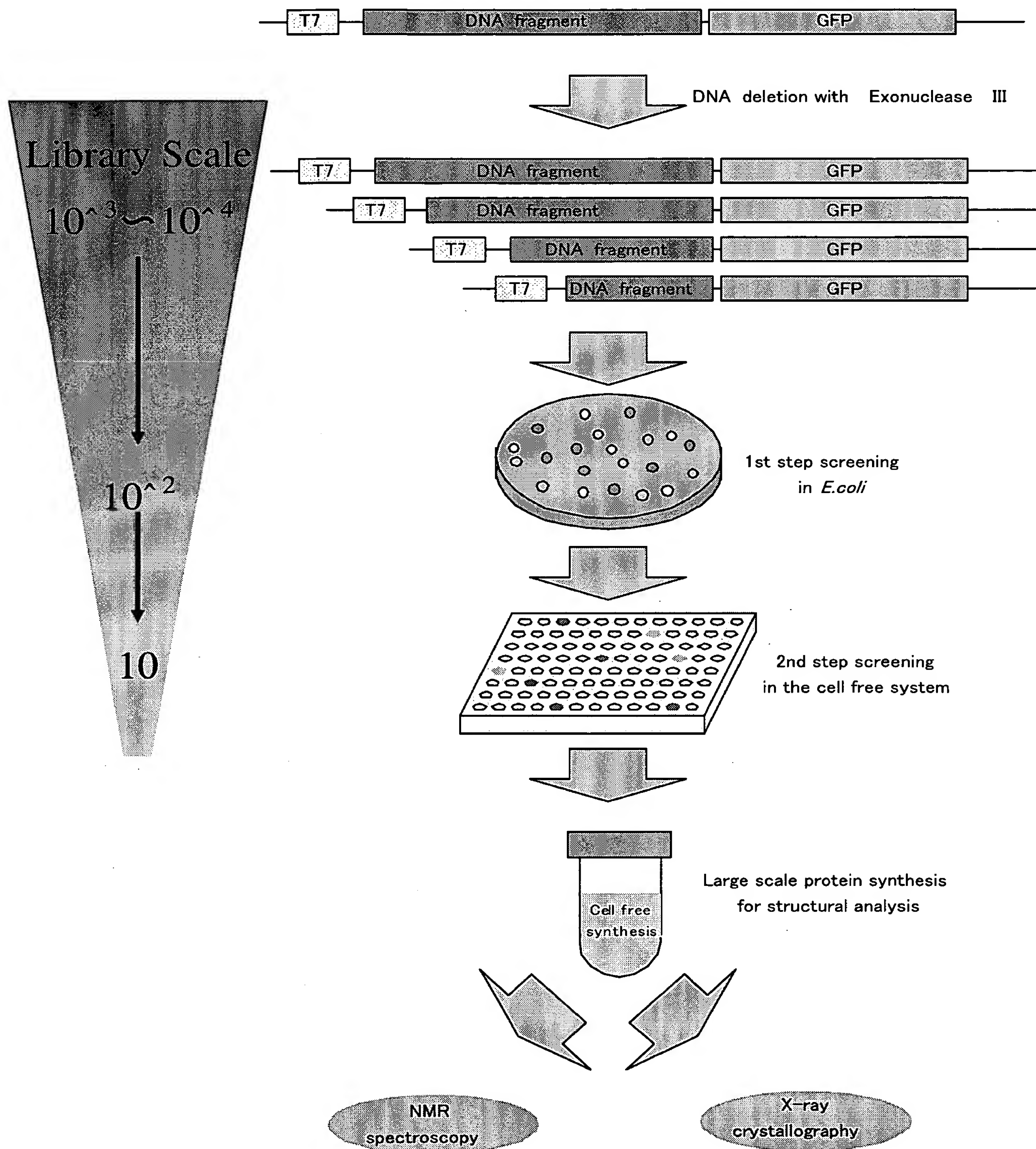
大腸菌での1段階目スクリーニング



無細胞合成での 2段階目スクリーニング



タンパク質ドメインの スクリーニングシステムの流れ

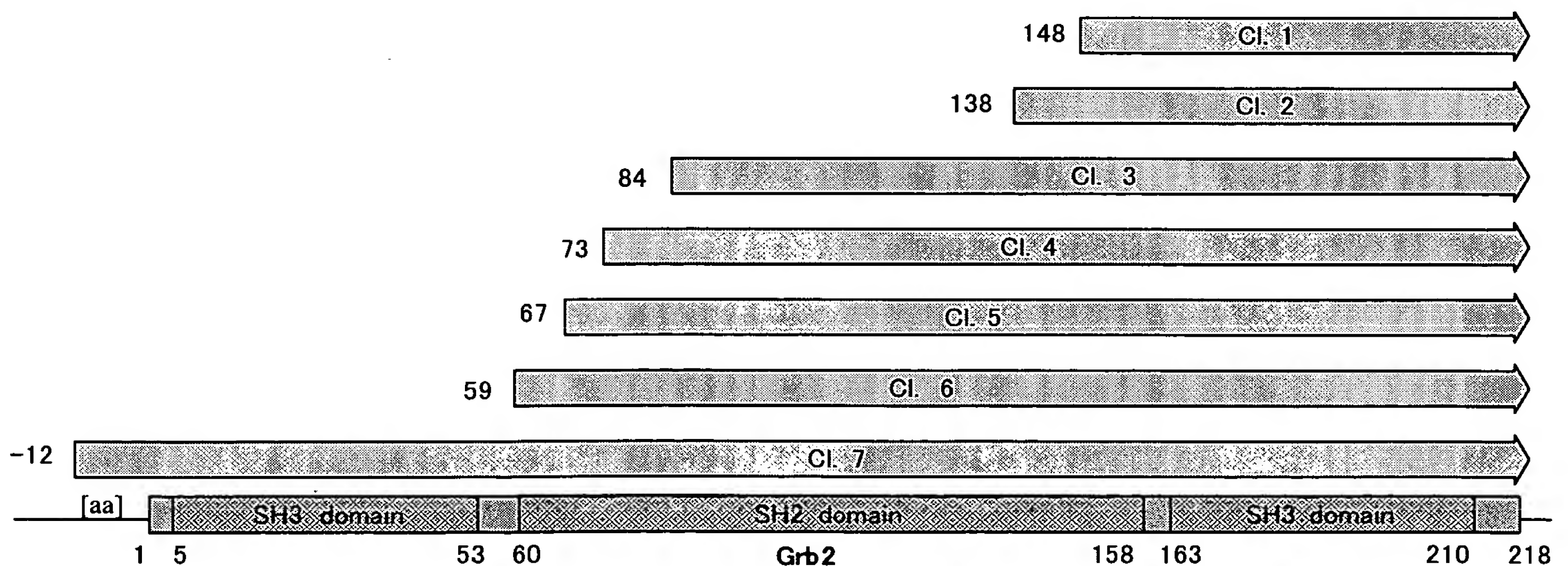


Grb2のドメインスクリーニング

• Grb2

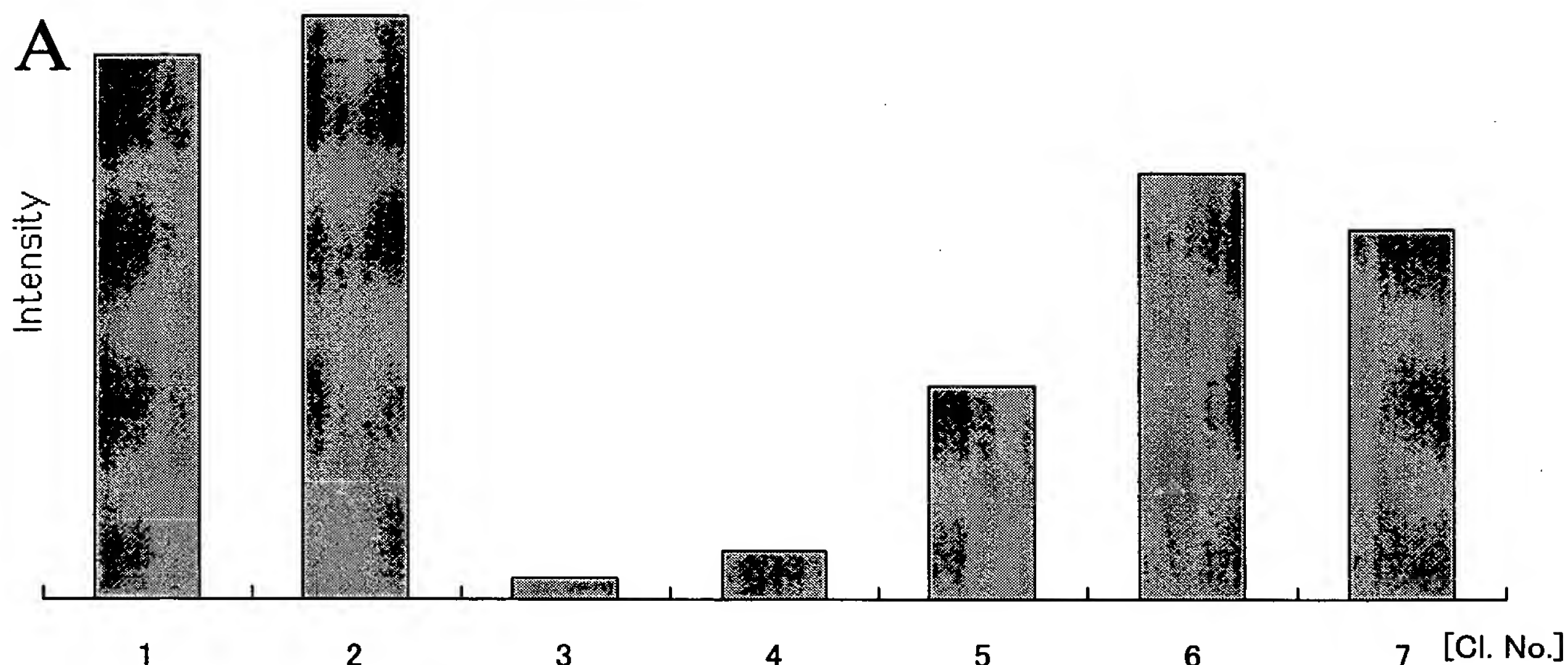
- 全長で可溶性 GFP-fusionは発光
 - SH3 SH2 SH3ドメインを含む
 - 立体構造がすでに解析されている
- (Maignan, S. et. Al. Science 268, 291-293 1995)

スクリーニングにより選択された Deleted cloneの位置

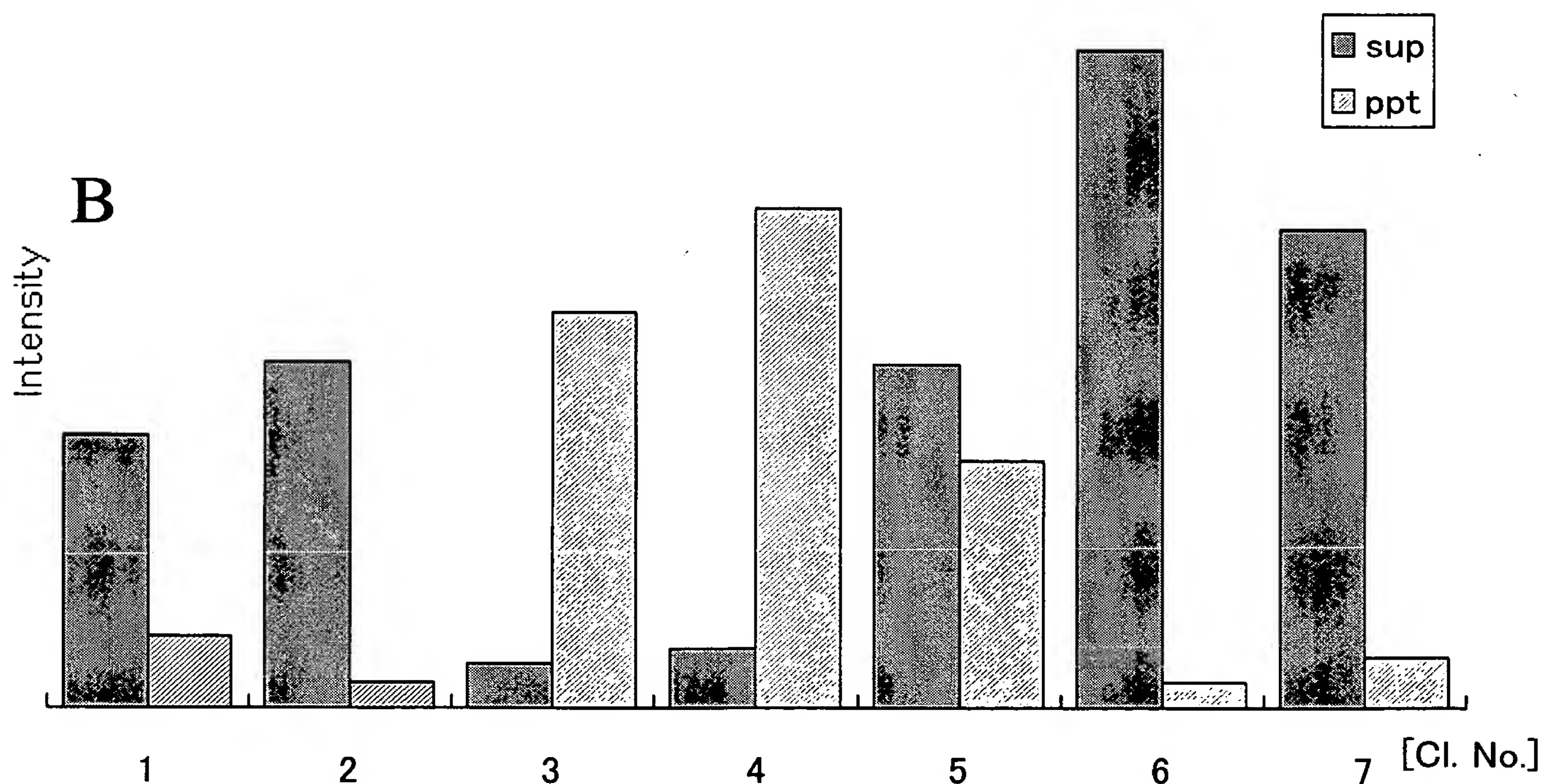


*ここでのドメインはアミノ酸配列の保存性から導き出された位置を示している

Grb2のDeleted cloneの解析(1)



A 無細胞合成による GFP-fusionの蛍光強度
(励起 485nm 発光 535nm)



B 無細胞合成による単体での可溶性
(^{14}C , scintillation counter)

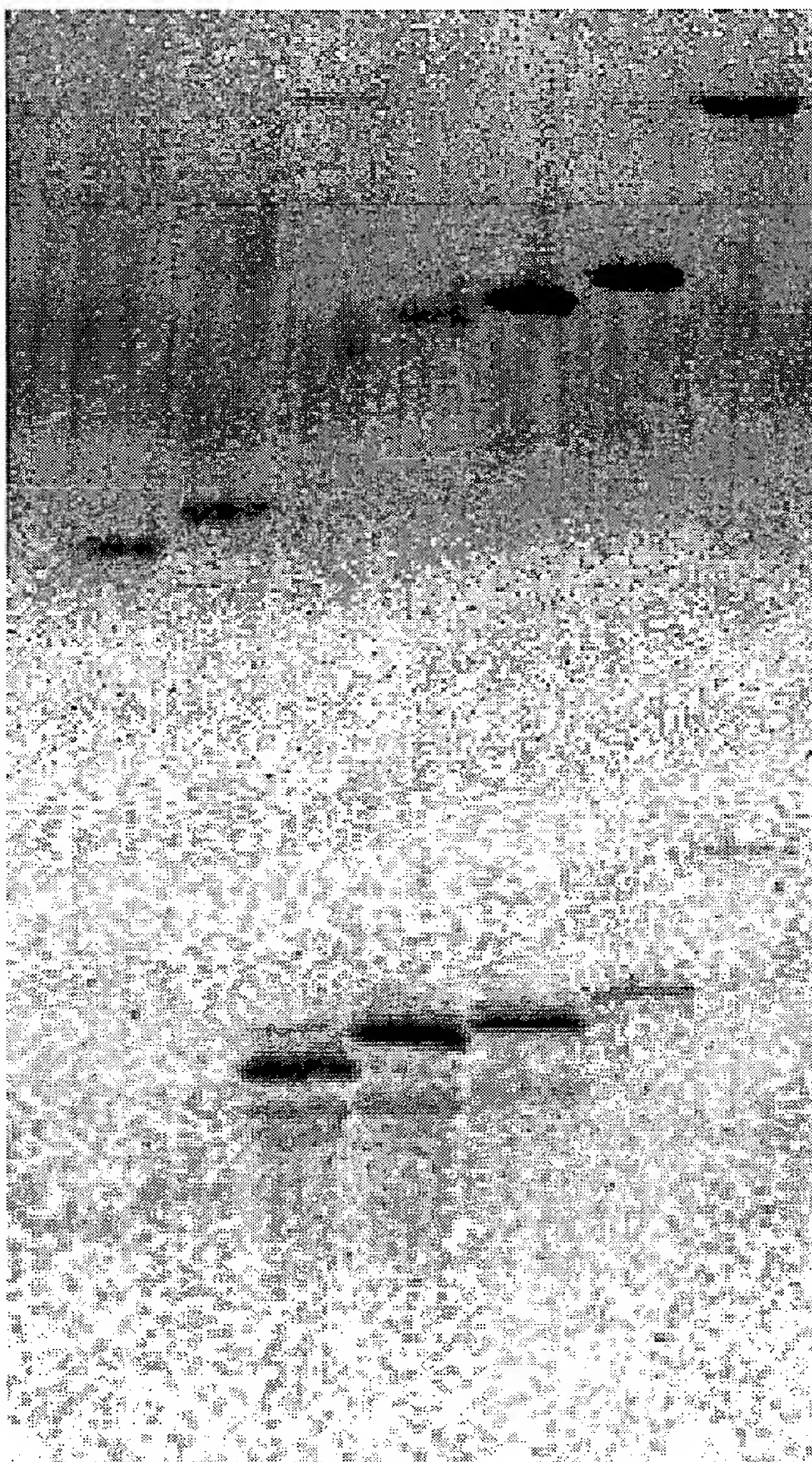
Grb2のDeleted cloneの解析(2)

C

1 2 3 4 5 6 7

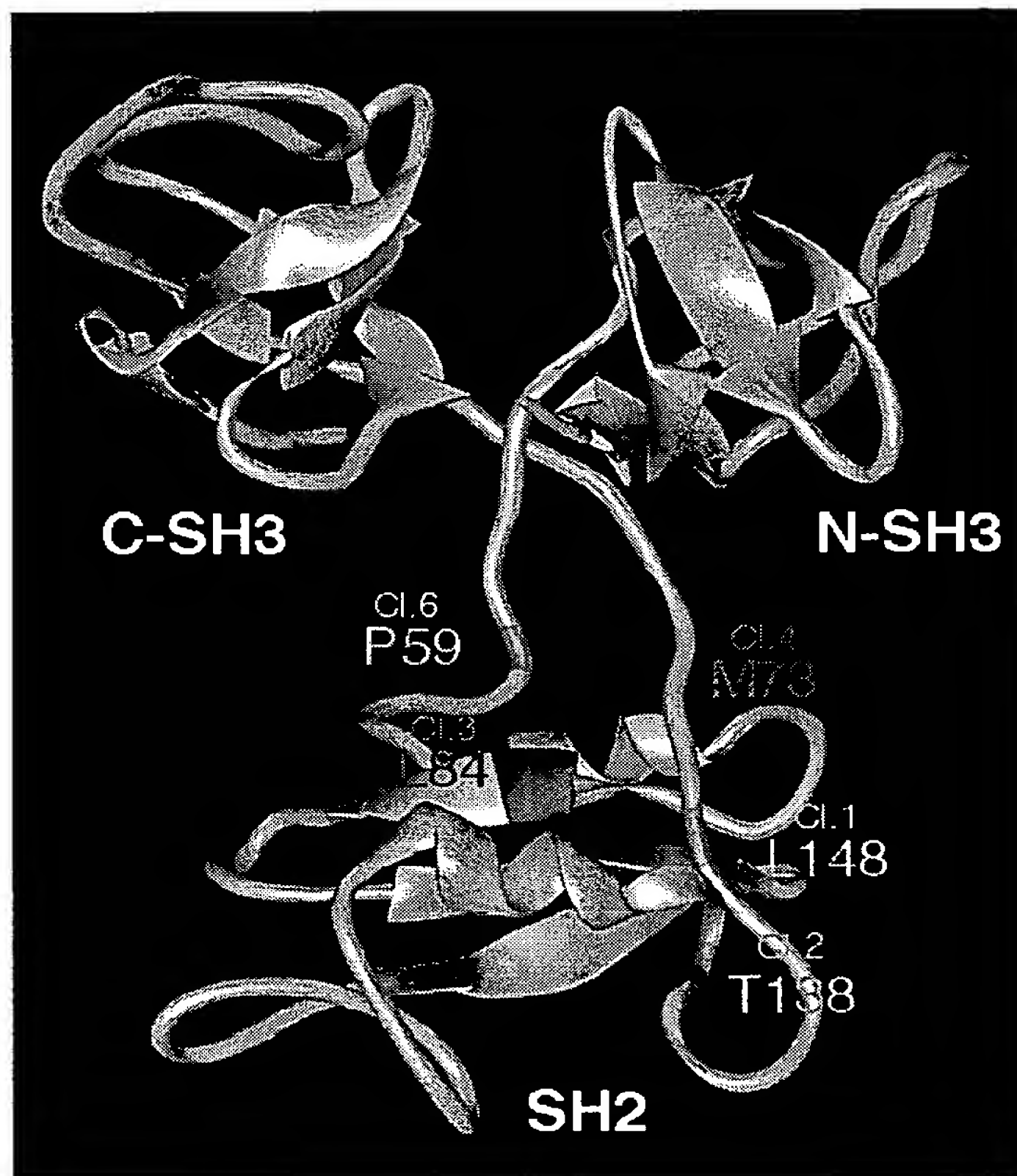
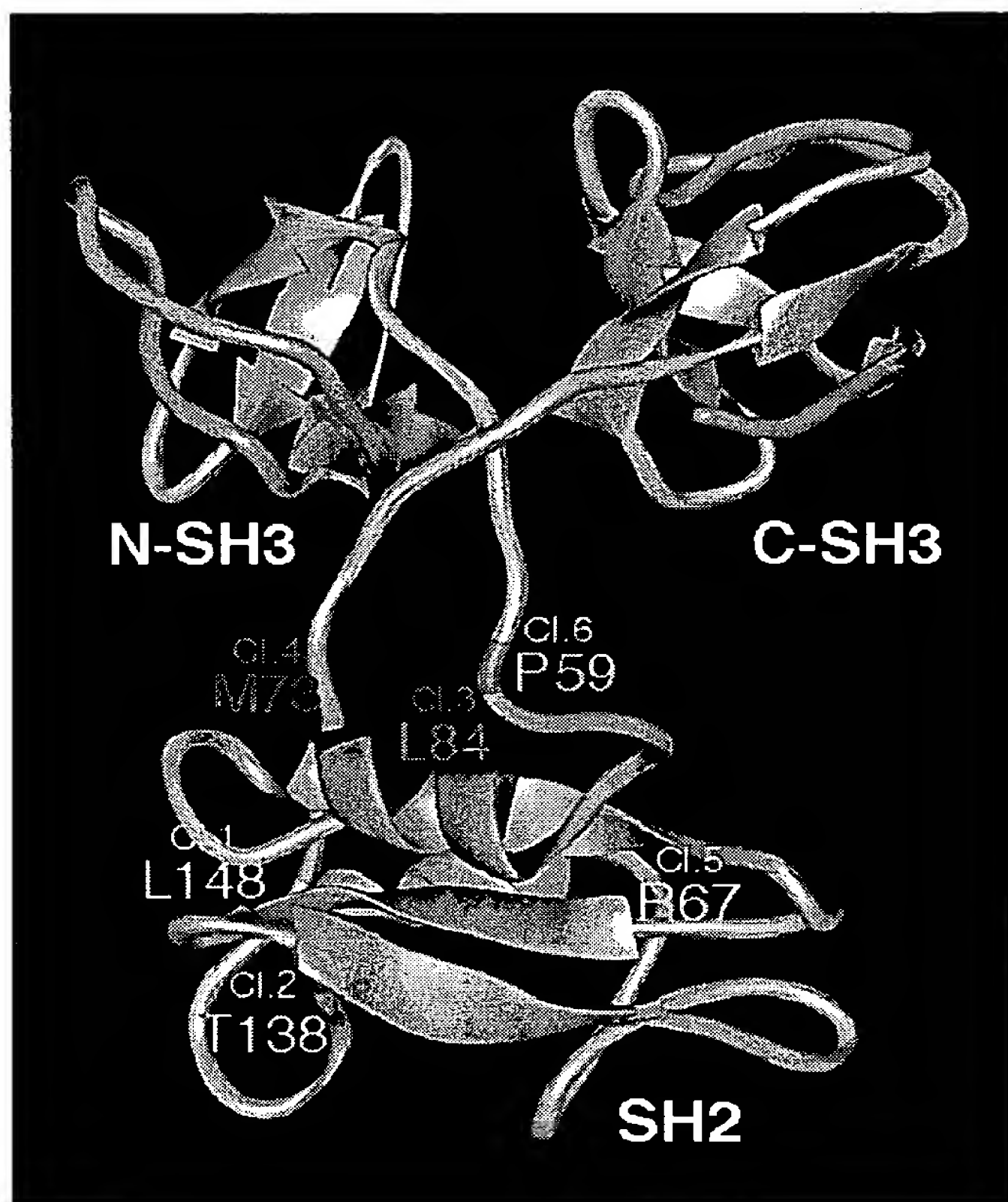
sup.

ppt.



C. 無細胞合成による単体のSDS-PAGE
(¹⁴C, MacBas)

立体構造とDeletion部位の比較



■ : Soluble

■ : Partially soluble

■ : Insoluble

結論

- Exonuclease IIIを用いることで、均一でサンプリング範囲の広いDNA randomly deleted library を作成することが可能になった。
- DNA deleted libraryとタンパク質C末端側でのGFP-fusionを組み合わせ、さらに、大腸菌と無細胞系を用いて2段階でスクリーニングすることにより、可溶性の高いドメインを、効率よく選び出すことが可能になった。
- Grb2においては、本系が導き出した可溶性の高いfragmentの境界と、立体構造上のドメイン境界との間に強い相関があった。

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☒ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.